Supplementary information

The supplementary information includes Table S1, Figure S1, Figure S2, Figure S3 and Figure S4.

Approach	Best F	Best P	Best R	AUC
This paper	.72	.68	.77	.64
mPb	.72	.68	.76	.64

Table S1: Comparison with *mPb*, using the boundary annotations as ground-truth. P and R are precision and recall, respectively; AUC is the area under the precision-recall curve.

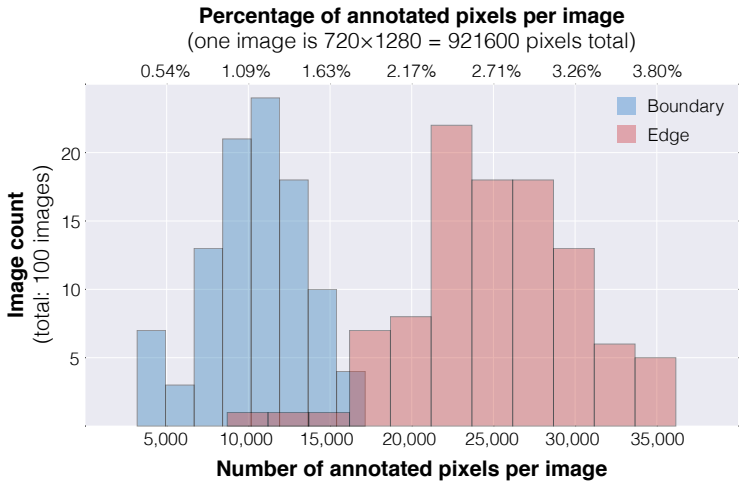

Figure S1: Histograms of the number of pixels labeled as contours, for either object boundaries (in blue) or edges (in red), averaged over all annotators and across the entire database.

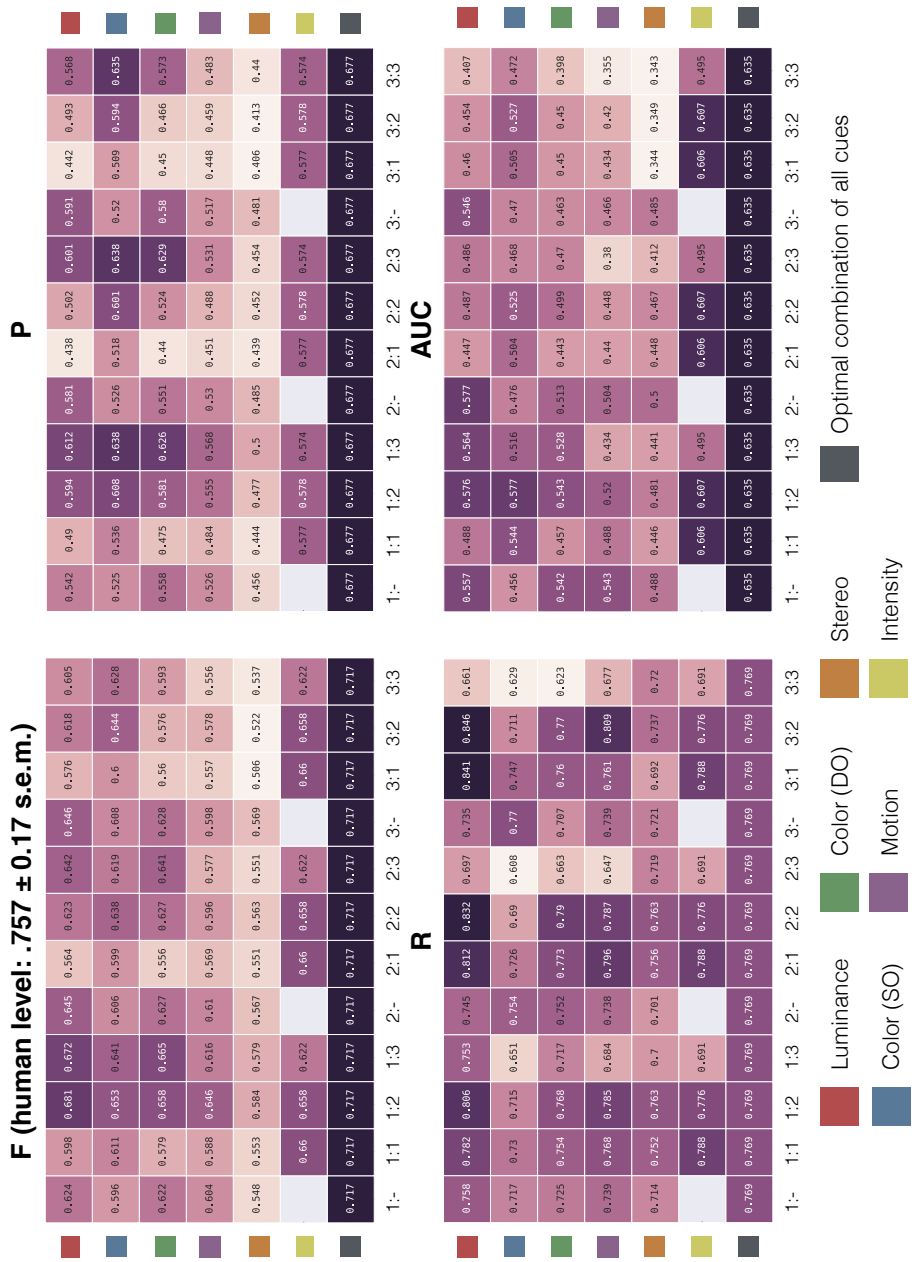

Figure S2: Cue accuracy measured as F score, precision-recall (P) and recall (R) as well as area under the precision-recall curve (AUC). Columns are labeled i:j where i stands for filter size and j for χ^2 operator size (a value of '-' means N/A). Note that the intensity cue is based on pixel intensities instead of filter responses, thus its score is defined only across 3 χ^2 sizes (values are replicated across filter sizes for easier comparison). The 'all' cue stands for the optimal combination of all cues for all filter sizes and χ^2 sizes. Its single score is replicated across filter and χ^2 operator sizes for easier comparison.

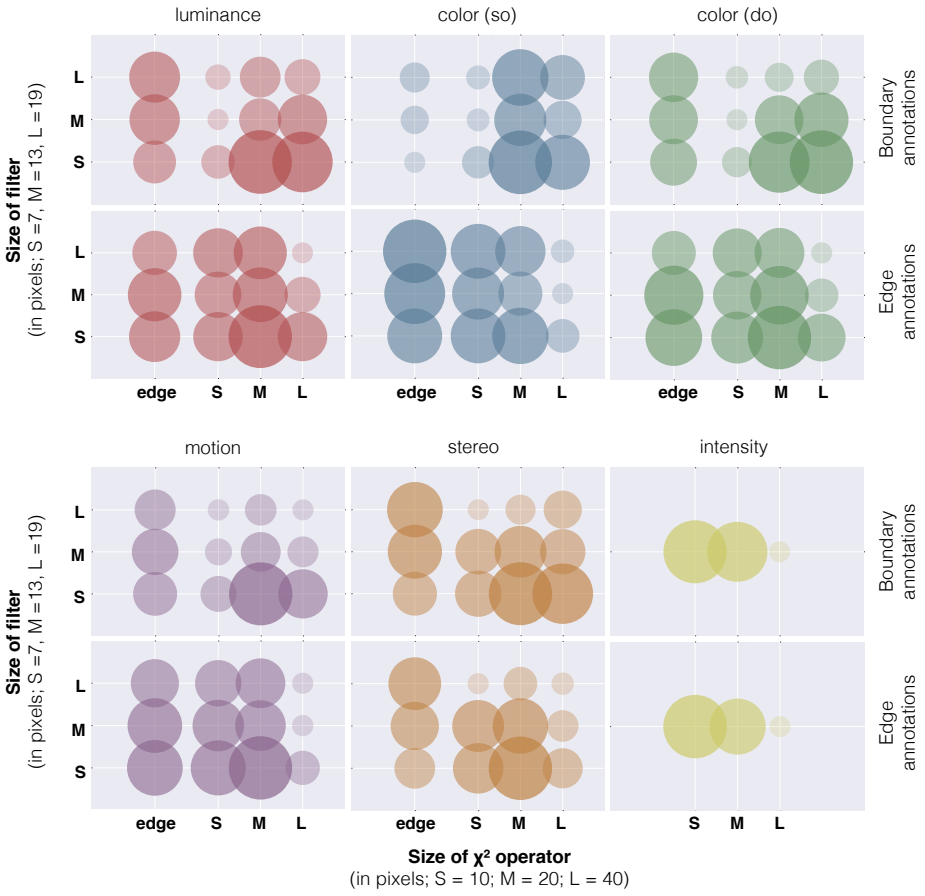

Figure S3: Cue accuracy (normalized F-score) for (high-level) boundary annotations (upper plots, "boundary annotations") and for (low-level) edge annotations (lower plots, "edge annotations"), across filter sizes (y-axis) and χ^2 operator sizes (x-axis). Disk transparency and radius indicate the F score of the corresponding classifier; we normalized disk sizes inside each plot. The region-based approach corresponds to χ^2 operator size= S (small), M (medium), L (large) (the actual values are given in the paper) and edge-based approach to "edge".

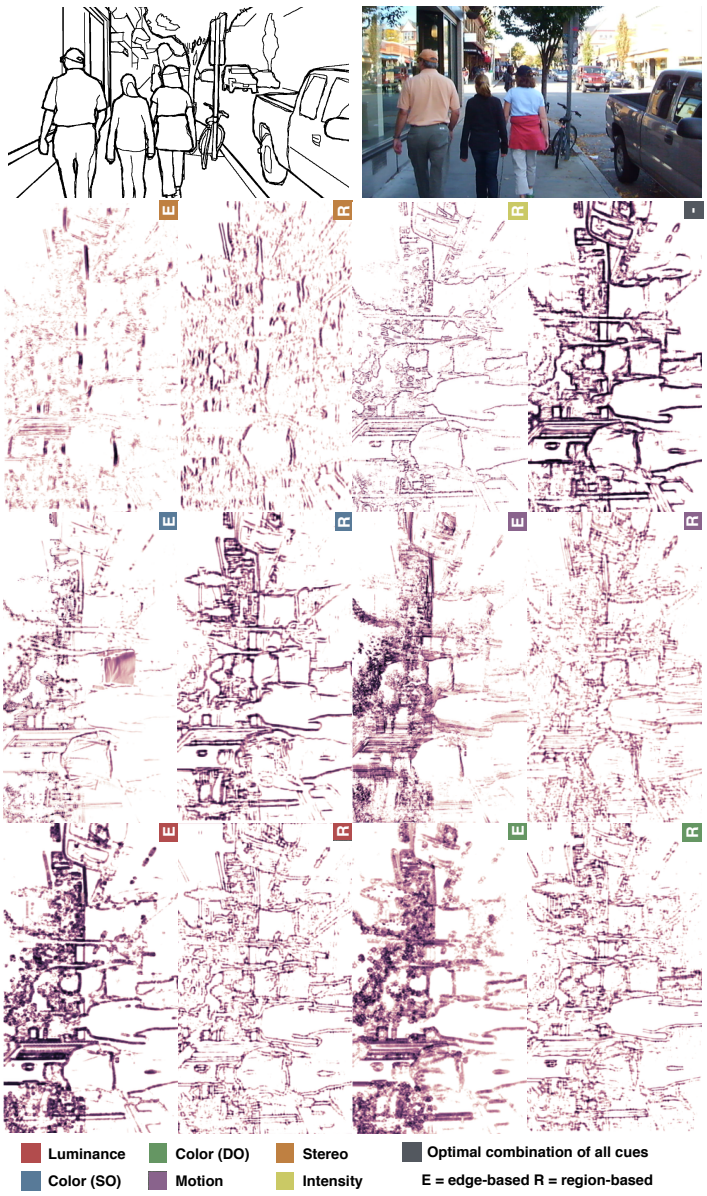

Figure S4: Boundary predictions (probability of boundary) from each cue, showing the edge-based *vs.* region-based approach (when applicable), computed on one of the sample scenes in Figures 1 and 2. Each output was thresholded at the optimal value that yielded the F score reported in the paper. For each cue and approach we show the output corresponding to the best-performing classifier across filter sizes and/or χ^2 sizes. We show ground-truth annotations (after a 5 pixel-wide dilation for visualization purposes) and original scene on top.